

Plataforma de Gestión y Operacionalización de Modelos de Aprendizaje Automático para la Distribución de Bienes de Consumo

Est. Dante Emanuel Tellez Guevara
Depto. Informática
Universidad de America de Europa
Cancún, México
dantetellez@yahoo.com

Resumen— Este estudio se centra en el diseño de una plataforma que operacionaliza modelos de aprendizaje automático (ML) en la distribución de bienes de consumo, una tarea clave para la integración efectiva de la inteligencia artificial (IA) en los entornos productivos empresariales. La investigación abarca metodologías y marcos que se adaptan a la madurez tecnológica, estructura organizacional, tamaño del equipo y cultura corporativa de una empresa hipotética. Este trabajo ofrece un enfoque integral que considera tanto la eficiencia técnica de los modelos de ML como su alineación con las estrategias y objetivos empresariales. Además, como ejemplo, se enfoca en tareas importantes de las empresas de distribución de bienes de consumo, tales como la predicción del tiempo de transporte, y la optimización de rutas de distribución. Esta plataforma no solo aporta conocimientos teóricos y metodológicos al campo académico, sino que también proporciona un marco práctico y adaptable para la integración eficiente de soluciones de IA en las empresas, asegurando su aplicabilidad y relevancia en el entorno empresarial actual y futuro.

Palabras Clave— Aprendizaje Automático, Operacionalización de modelos de aprendizaje automático, Distribución de Bienes de Consumo, Modelado Predictivo, Optimización de rutas.

Abstract— This study focuses on the design of a platform that operationalizes machine learning (ML) models in consumer goods distribution, a crucial task for the effective integration of artificial intelligence (AI) in business production environments. The research covers methodologies and frameworks that adapt to the technological maturity, organizational structure, team size, and corporate culture of a hypothetical company. This work offers a comprehensive approach that considers both the technical efficiency of ML models and their alignment with business strategies and goals. Additionally, as an example, it focuses on important tasks for consumer goods distribution companies, such as predicting transportation time and optimizing distribution routes. This platform not only contributes theoretical and methodological insights to the academic field but also provides a practical and adaptable framework for the efficient integration of AI solutions in businesses, ensuring their applicability and relevance in the current and future business environment.

Keywords— Machine Learning, Deployment of Machine Learning Models, Consumer Goods Distribution, Predictive Modeling, Route Optimization.

I. INTRODUCCIÓN

Actualmente, dada la importancia de la transformación digital, la industria busca optimizar y actualizar sus procesos con el objetivo de mantenerse competitiva. El sector de distribución de bienes de consumo, que abarca desde la producción hasta la entrega final de los productos a los consumidores, no es la excepción. Este sector desempeña un papel crucial al garantizar que una variedad de productos, como alimentos, electrodomésticos y ropa, lleguen de manera eficiente a los puntos de venta y, en última instancia, a los consumidores [1].

Diferentes trabajos se enfocan en la correcta gestión de la distribución de productos de consumo. Generalmente, estos sistemas de gestión involucran modelos como Six Sigma, Malcolm Baldrige o el ciclo PDCA [2]. Otros enfoques involucran una metodología más experimental involucrado estudios cuantitativos [3]. Existen también trabajos que integran modelos de Aprendizaje de Máquina (ML, por sus siglas en inglés) para automatizar y eficientizar tareas. Por ejemplo, se han propuesto modelos para la gestión de la flota vehicular [4], modelos para la mejora de rutas [5], modelos para el análisis de costos y tiempos de entrega [6], entre otros.

La principal contribución de este trabajo está en el diseño de una plataforma integral de gestión y operacionalización de modelos de aprendizaje automático para la distribución de bienes de consumo. Se pretende no únicamente evaluar modelos de ML para el sector de distribución, sino estudiar de manera integral el despliegue de estos modelos, y su alineación con las estrategias de negocio. La plataforma propuesta es una solución tecnológica integral diseñada para implementar y optimizar modelos de aprendizaje automático en el ámbito de la distribución. Facilita la integración de estos modelos en los procesos operativos, gestionando datos, y optimizando rutas de distribución y tiempos de transporte. Adaptada a la estructura y cultura corporativa, la plataforma ofrece herramientas para monitorear y evaluar el rendimiento de los modelos, apoyar decisiones estratégicas y mejorar la eficiencia logística, transformando los avances en inteligencia artificial en soluciones prácticas y efectivas.

El resto del artículo está organizado de la siguiente forma. La Sección II describe la metodología y los conceptos clave para el desarrollo del proyecto. La Sección III detalla los modelos de ML y gestión utilizados. Después, la Sección IV muestra los resultados. Por último, la Sección V presenta las conclusiones.



Fig. 1. Tareas genéricas del sector de la distribución de bienes de consumo. Para la plataforma propuesta, se evalúan modelos para i) La predicción del tiempo de transporte; y ii) La optimización de rutas de distribución.

II. METODOLOGÍA

Para el diseño de la plataforma de gestión, el presente proyecto aborda dos tareas: i) La tarea de entrenar modelos de ML con el potencial de mejorar el sector de la distribución de bienes de consumo, y ii) La tarea de evaluar estrategias de integración y operacionalización de estos modelos, pues se considera importante no sólo entrenar modelos sino integrarlos adecuadamente a la estrategia del negocio, además de integrarlos en una plataforma única. La (Fig. 1) muestra las tareas principales que generalmente ejecutan las empresas de distribución de bienes de consumo. Particularmente, los modelos explorados en este proyecto se enfocan en la predicción del tiempo de transporte; y la optimización de rutas de distribución. Estos casos de uso fueron elegidos como ejemplo para la operacionalización e integración con la plataforma de gestión del sector de distribución de productos.

A continuación, se describe con más detalle los componentes de la plataforma propuesta.

A. Plataforma de Gestión y Operacionalización

Como se ha mencionado, el verdadero valor de los modelos se alcanza cuando se operacionalizan, es decir, se despliegan en producción. La operacionalización permite que el negocio utilice los modelos en sus operaciones diarias. Para lograr una implementación exitosa y que los modelos sean efectivamente utilizados en producción, es necesario abordar varios aspectos críticos. En la (Fig. 2) se presentan tres etapas clave para el despliegue en producción.

En el marco de operacionalización propuesto, primero se realiza una evaluación exhaustiva que considera la madurez tecnológica, la gestión de datos, la cultura organizacional, las habilidades del personal, y los requisitos de seguridad y cumplimiento. Luego, se desarrolla una estrategia sólida que identifica las fuentes de datos esenciales, define la arquitectura de software y establece una hoja de ruta de Inteligencia Artificial (IA, por sus siglas en inglés) que se alinea con los objetivos del negocio. La etapa final se centra en la acción; se desarrollan y ajustan los modelos de IA, se integran en los sistemas existentes y se establecen protocolos para su monitoreo continuo, mantenimiento y optimización escalable. Este proyecto pone especial énfasis en los pasos necesarios para una implementación exitosa y una adaptación efectiva de los modelos en entornos de producción reales.

Para la operacionalización de los modelos, se sigue la metodología propuesta en la (Fig. 2). Descrita con mayor detalle a continuación.



Fig. 2. Marco para la operacionalización de modelos de ML. Se inicia con la Evaluación, que examina la preparación tecnológica y organizacional. Luego, la Estrategia define el plan de datos y software para alinear la IA con los objetivos empresariales. Finalmente, la etapa de Acción se enfoca en el desarrollo y mantenimiento continuo de los modelos de IA.

Evaluación.

Madurez Tecnológica: En la industria de distribución de bienes de consumo, la madurez tecnológica es clave para el éxito. Los proyectos de predicción del tiempo de transporte, optimización de rutas requieren sistemas tecnológicos avanzados que puedan manejar grandes volúmenes de datos y ejecutar algoritmos de aprendizaje automático de manera eficiente. Mantenerse al día con las últimas tendencias tecnológicas es crucial para asegurar la competitividad.

Gobierno de Datos: La gestión adecuada de los datos es esencial para la precisión de los modelos. Para los proyectos de predicción y optimización, se necesita un sólido gobierno de datos, lo que implica establecer políticas claras para la recopilación, almacenamiento y uso seguro de la información. Asegurar la calidad de los datos y cumplir con las regulaciones de protección de datos son aspectos fundamentales.

Cultura Organizativa: La implementación exitosa de proyectos en la distribución de bienes de consumo requiere un cambio cultural en la organización. Fomentar una cultura de innovación, aprendizaje continuo y apertura a nuevas tecnologías es crucial. La colaboración entre los equipos de logística, tecnología y datos juega un papel vital en el éxito de estos proyectos.

Talento y Habilidades: Tener el personal adecuado con las habilidades necesarias es esencial. La industria necesita expertos en análisis de datos, aprendizaje automático, programación y logística. Los proyectos se benefician de equipos multidisciplinarios que combinen conocimientos técnicos con experiencia en la industria. Invertir en la formación y desarrollo del personal es clave.

Seguridad y Cumplimiento Normativo: La protección de datos y el cumplimiento de regulaciones son cruciales. Los proyectos deben abordar la seguridad de la información y cumplir con regulaciones específicas de la industria, como la protección de datos y la seguridad de la cadena de suministro. Implementar medidas de seguridad y gestionar el cumplimiento normativo son aspectos fundamentales.

Estrategia

Fuentes de Datos: Identificar y gestionar las fuentes de datos es esencial para los proyectos en la distribución de bienes de consumo. Para la predicción del tiempo de transporte y la optimización de rutas, es crucial recopilar datos relevantes, como seguimiento de pedidos, información de tráfico, condiciones climáticas y datos de inventario.

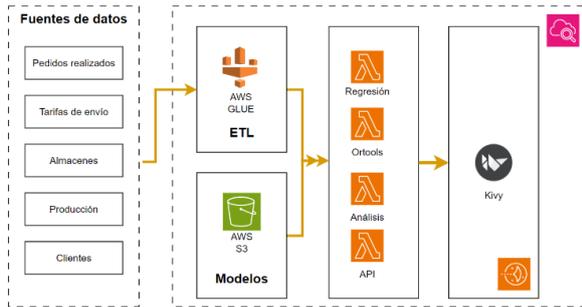


Fig. 3. Arquitectura propuesta en AWS para la gestión y operacionalización de modelos de ML en distribución de bienes de consumo.

La estrategia debe asegurar la adquisición en tiempo real y la integración efectiva de los datos para mantener modelos actualizados y precisos. El conjunto de datos utilizado consiste en varias hojas de un archivo de Excel, cada una conteniendo información específica relacionada con la logística de la cadena de suministro [7].

Arquitectura de Software: La arquitectura de software debe ser escalable y flexible para manejar grandes volúmenes de datos y ejecutar algoritmos de aprendizaje automático eficientemente. Debe integrarse con los sistemas existentes y permitir futuras expansiones. Elegir herramientas y tecnologías adecuadas es clave para una implementación exitosa. Considerando una arquitectura inicial con capacidades de escalamiento se propone para cada tarea crítica lo siguiente mostrado también en la (Fig. 3):

- **Ingesta de Datos:** Los datos se recopilan y almacenan en Amazon S3. AWS Glue realiza tareas de ETL para preparar los datos.
- **Entrenamiento de Modelos:** Los datos preparados se utilizan para entrenar modelos con scripts de Python, y los modelos se cargan en un S3.
- **Despliegue y Acceso:** Los modelos entrenados se despliegan a producción mediante Amazon Lambdas y se exponen a través de Amazon API Gateway.
- **Monitoreo:** Amazon CloudWatch supervisa el rendimiento de los modelos y la plataforma en general, generando alertas y realizando ajustes según sea necesario.
- **Vista:** Se usa Kivy para la interacción con la plataforma. Kivy está en una instancia de Linux de AWS Lightsail.

Estrategia de IA y Alineación de Negocios: La estrategia de inteligencia artificial (IA) debe estar alineada con los objetivos comerciales, como la reducción de costos y la mejora de la eficiencia. Debe definir claramente cómo los modelos de IA contribuirán a estos objetivos y cómo se medirá su impacto en el negocio.

Roadmap de IA: Establecer un roadmap de IA es esencial para planificar la implementación de proyectos a lo largo del tiempo. Este debe incluir hitos y entregables clave, como desarrollo, pruebas y despliegue de modelos, así como monitorización y ajustes continuos. Considerar la escalabilidad de las soluciones de IA a medida que la empresa crece es fundamental.

Acción

Desarrollo de Modelos de IA: Desarrollar modelos de IA robustos es crucial. Esto incluye seleccionar algoritmos adecuados, preparar y entrenar datos, y evaluar los modelos. Los modelos deben generar predicciones precisas que faciliten la toma de decisiones en la distribución de bienes de consumo.

Modelo I. Predicción del tiempo de transporte

En el contexto del tiempo de transporte, los modelos buscan estimar la duración que tomará el envío de bienes de un punto a otro. El tiempo de transporte, que es un factor crucial en la logística y planificación de envíos, puede verse afectado por una variedad de elementos. Comprender estos factores es esencial para la creación de modelos predictivos eficientes y precisos en este ámbito [8, 9]

Para predecir el tiempo de transporte, se pueden utilizar diversos métodos de regresión como la Regresión Lineal [10], y No Lineal [11], los Modelos Basados en Árboles, como Random Forest [12], y Gradient Boosting [13], y los modelos de aprendizaje profundo como las Redes Neuronales [14]. Cada uno de estos métodos tiene sus propias fortalezas y se seleccionan en función de la naturaleza de los datos y la relación entre las variables de entrada y el tiempo de transporte [15, 16].

Se puede modelar una regresión para la predicción del tiempo de transporte de la siguiente forma

$$\hat{y} = f(X). \quad (1)$$

Donde \hat{y} es el tiempo de transporte predicho, y $X = [x_1, x_2, \dots, x_n]$ representan los diversos factores o características que influyen en el tiempo de transporte, como la distancia, las condiciones del tráfico, el clima, entre otras.

El desarrollo de un modelo predictivo comienza con la recolección y el procesamiento de datos relevantes. Estos datos son fundamentales para la construcción del modelo, ya que su calidad y precisión determinan la efectividad de las predicciones. Una vez que los datos están listos, se selecciona un modelo adecuado basado en las características de los datos y el tipo de predicción deseada. Este modelo se entrena luego con los datos recopilados, ajustándolo para reflejar las relaciones y patrones inherentes a los datos.

La validación del modelo es un paso crucial en este proceso. Una técnica comúnmente utilizada es la validación cruzada [15], donde el conjunto de datos se divide en partes, y el modelo se entrena y prueba en diferentes segmentos de estos datos. Esto ayuda a garantizar que el modelo sea capaz de generalizar y no solo ajustarse específicamente a los datos con los que fue entrenado.

La precisión y utilidad del modelo se evalúan mediante métricas específicas. Dos de las más importantes son el Error Cuadrático Medio (MSE) [17]. El MSE se calcula como:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2)$$

Donde y_i es el valor real y \hat{y}_i es el valor predicho por el modelo. Luego de completar el proceso de validación, se procede a afinar el modelo según sea necesario, con el objetivo de optimizar su rendimiento antes de su despliegue para efectuar predicciones en situaciones de la vida real. Esta fase de implementación adquiere una relevancia crítica en sectores como la logística y la planificación de rutas, donde la exactitud en las predicciones del tiempo de transporte se convierte en un

factor vital para la toma de decisiones estratégicas. Además, dentro del ámbito de interés de este proyecto de tesis, también se aborda el Modelo de Ruta Óptima, el cual será detallado a continuación.

Modelo II. Optimización de rutas

La optimización de rutas es un área clave en la planificación logística, donde el objetivo es minimizar el tiempo total de transporte entre varios puntos [18, 19]. El objetivo fundamental es determinar la trayectoria que minimiza el tiempo total de transporte [20]. Para desarrollar un modelo de Ruta Óptima, se toman en consideración las siguientes suposiciones:

- Se parte del supuesto de que los tiempos de transporte entre los distintos puntos son conocidos y representativos, lo que permite la creación de una matriz que contiene información sobre las distancias o tiempos necesarios para trasladarse de un lugar a otro.
- Se simplifica el enfoque al omitir la consideración de elementos específicos de los pedidos, centrándose exclusivamente en la determinación de las rutas más eficientes entre los puntos de origen y destino.

Para abordar este desafío, es posible emplear técnicas de programación lineal o recurrir a algoritmos especializados en la optimización de rutas, como el algoritmo de Dijkstra diseñado para grafos dirigidos [21].

$$\text{Minimizar } Z = \sum_{i=1}^n \sum_{j=1, j \neq i}^n t_{ij} \cdot x_{ij}, \quad (4)$$

donde t_{ij} representa el tiempo de transporte del punto i al punto j . x_{ij} es una variable binaria que es 1 si se elige la ruta de i a j y 0 en caso contrario, y n es el número total de puntos en la red. Las restricciones típicas en un modelo de ruta óptima incluyen: i) Cada punto solo puede ser visitado una vez, y ii) La ruta debe ser conexa y volver al punto de origen.

$$\sum_{i=1, i \neq j}^n x_{ij} = 1, \text{ para todo } j. \quad (5)$$

$$\sum_{j=1, j \neq i}^n x_{ij} = 1, \text{ para todo } i. \quad (6)$$

Integración y Despliegue de Software: La integración de los modelos de IA en los sistemas existentes es un paso crítico. Esto puede implicar el uso de APIs o la integración directa en el software. Un despliegue planificado y controlado garantiza una transición fluida a la operación en tiempo real.

Monitoreo y Mantenimiento de Modelos: Una vez en producción, es vital monitorear y mantener los modelos continuamente. Esto incluye supervisar su precisión y rendimiento, establecer alertas para posibles desviaciones y realizar mantenimientos proactivos para actualizar los modelos conforme cambian los datos y las condiciones.

Escala y Optimización: A medida que los proyectos ganan tracción, asegurar que los sistemas puedan escalar para manejar más datos y transacciones es crucial. La optimización continua de los modelos y procesos garantiza la eficiencia y calidad de las soluciones de IA.

III. PLATAFORMA DE GESTIÓN Y OPERACIONALIZACIÓN

En la (Fig. 4) se muestra el prototipo de la plataforma.

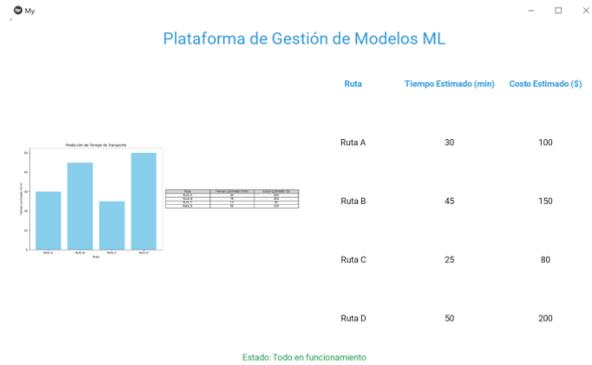


Fig. 4. Prototipo de la plataforma.

Cuando se inicia la aplicación, se muestra un título que dice "Plataforma de Gestión de Modelos ML". La interfaz está organizada verticalmente para facilitar la navegación. Primero, se presenta una imagen que representa la predicción de transporte. Esta imagen ocupa la parte izquierda de la pantalla y muestra un gráfico relacionado con la optimización de rutas o el análisis de datos de transporte.

A la derecha de la imagen, se muestra una tabla con información sobre rutas óptimas. La tabla incluye columnas para la "Ruta", "Tiempo Estimado (min)" y "Costo Estimado (\$)", permitiendo al usuario comparar fácilmente diferentes opciones de rutas.

En la parte inferior, hay una sección de estado que informa al usuario sobre el estado general de la plataforma, mostrando un mensaje como "Estado: Todo en funcionamiento". Esto le permite al usuario ver rápidamente si la plataforma está funcionando correctamente.

La interfaz, con un fondo blanco y elementos visuales claros, está diseñada para ejecutarse en una instancia de Linux en AWS Lightsail, asegurando que sea accesible y se integre con otros componentes de la arquitectura en la nube. La combinación de gráficos y tablas permite al usuario interactuar directamente con los resultados de los modelos y las predicciones, haciendo de la plataforma una herramienta útil para la gestión de modelos de Machine Learning.

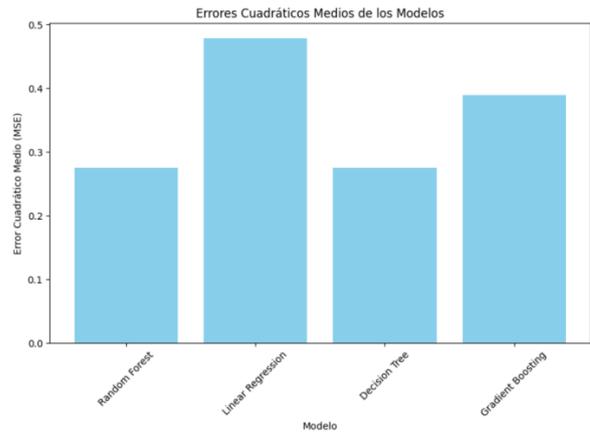


Fig. 5. Comparación de error MSE por algoritmo.

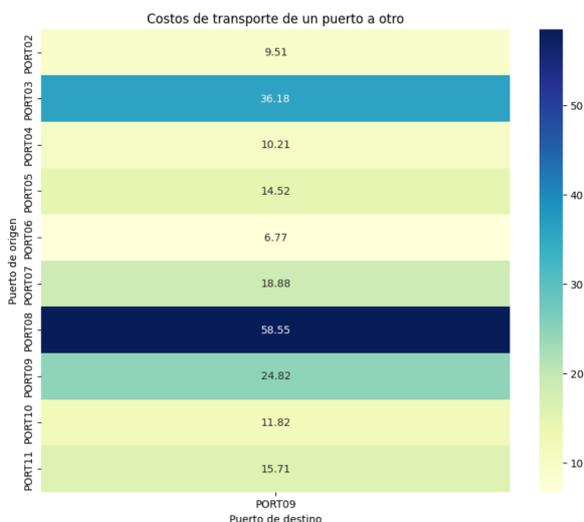


Fig. 6. Costo de la ruta de cada puerto al puerto 09. Estos valores proporcionan una visión general de los costos de transporte desde diferentes puertos de origen hacia el destino 'PORT09'.

IV. RESULTADOS

A. Predicción de los tiempos de transporte

Los resultados de los modelos se ilustran en las figuras correspondientes. El modelo Random Forest obtuvo un error cuadrático medio (MSE) de aproximadamente 0.275. Una de sus principales ventajas es su capacidad para manejar grandes conjuntos de datos con muchas características sin necesidad de ajustes extensivos. Al combinar múltiples árboles de decisión, este modelo captura relaciones no lineales y maneja la multicolinealidad de las características, lo que lo hace robusto frente al sobreajuste. Este rendimiento se debe a su habilidad para identificar relaciones complejas entre las características.

Por otro lado, el modelo de regresión lineal mostró un MSE más alto, alrededor de 0.478. Aunque este modelo es simple y fácil de interpretar, su rendimiento puede verse limitado cuando la relación entre las características y la variable objetivo no es lineal. Sin embargo, su principal ventaja radica en la facilidad con la que se puede comprender cómo cada característica influye en la predicción.

El modelo de árbol de decisión presentó un MSE similar al del Random Forest, cercano a 0.275. Los árboles de decisión son valiosos por su capacidad para manejar características numéricas y categóricas sin necesidad de preprocesamiento extenso. Además, son fáciles de interpretar y visualizar, facilitando la comprensión de las decisiones del modelo.

El modelo de Gradient Boosting mostró un MSE de aproximadamente 0.389. Este método es eficaz para mejorar la precisión del modelo al combinar múltiples árboles de decisión débiles, y es particularmente útil para datos heterogéneos y relaciones complejas. Sin embargo, puede ser más propenso al sobreajuste si los hiperparámetros no se ajustan adecuadamente. Aunque su rendimiento fue inferior al del Random Forest y el árbol de decisión en este conjunto de datos, el Gradient Boosting sigue siendo valioso por su capacidad de iterar y corregir errores para mejorar la precisión.

B. Optimización de rutas

Primero, se asume que los tiempos de transporte registrados en el DataFrame FreightRates son representativos y se utilizarán para construir una matriz de distancias o tiempos entre los puertos involucrados. Esta matriz es fundamental para calcular las rutas óptimas.

En segundo lugar, para simplificar el análisis y centrarse en la logística del transporte entre puertos, se ha decidido no considerar los detalles específicos de los pedidos individuales del DataFrame OrderList en este análisis. Esto permite una evaluación y optimización de las rutas entre los puntos de origen y destino sin complicar el modelo con las características detalladas de los pedidos.

El objetivo principal de este modelo es identificar las rutas que minimicen el tiempo total de transporte entre los puertos, aprovechando la información disponible en los DataFrames. Encontrar rutas eficientes puede reducir costos y aumentar la eficiencia en la cadena de suministro. Cabe mencionar que este modelo es una versión simplificada y que en una implementación real podrían ser necesarias consideraciones adicionales, como la capacidad de carga de los transportistas, restricciones de tiempo y condiciones de tráfico. Además, la selección del algoritmo para encontrar las rutas óptimas dependerá de la complejidad y las necesidades específicas de la cadena de suministro.

La (Fig. 4) ilustra los costos de transporte desde varios puertos de origen hacia el puerto de destino 'PORT09'. Cada fila representa un puerto de origen distinto, y la columna 'PORT09' muestra los costos asociados al transporte desde estos puertos hacia el destino. Por ejemplo, el costo de transporte desde 'PORT02' hacia 'PORT09' es de aproximadamente 9.51, mientras que desde 'PORT03' es de alrededor de 36.18, entre otros.

Estos valores proporcionan una visión general de los costos de transporte y son útiles para la planificación de rutas, la evaluación de costos logísticos y la toma de decisiones en la gestión de la cadena de suministro. La predominancia de ceros en la matriz sugiere que muchas ubicaciones no tienen conexiones directas, lo que podría indicar que la mayoría de las ubicaciones no están directamente vinculadas y, por lo tanto, el tiempo de viaje entre ellas es nulo. Solo un subconjunto de ubicaciones tiene conexiones directas, lo que explica los valores distintos de cero en esas posiciones específicas. Además, la alta proporción de ceros podría deberse a la disponibilidad limitada de información para calcular los tiempos de viaje entre ciertas ubicaciones.

C. Operacionalización de modelos

Se consideró una arquitectura en AWS, y la herramienta Kivy para la interfaz.

V. CONCLUSIONES

En este trabajo, se subraya la importancia de la implementación de modelos de aprendizaje automático en la distribución de bienes de consumo, destacando cómo la integración de la inteligencia artificial puede transformar los entornos empresariales de producción. Se examinan metodologías y frameworks que se adaptan a distintos niveles de desarrollo tecnológico y a la cultura organizacional de una empresa hipotética, abordando los desafíos organizacionales y técnicos que deben superarse para adoptar eficazmente la inteligencia artificial. Además de centrarse en la eficiencia

técnica de los modelos de ML, se enfatiza la necesidad de alinearlos con las estrategias y objetivos de la empresa.

Para abordar los aspectos clave de la distribución de bienes de consumo, se desarrollaron diversos modelos de aprendizaje automático. Estos modelos se enfocaron en la predicción del tiempo de transporte, y la optimización de rutas, utilizando técnicas de machine learning y programación lineal para optimizar sus respectivas funciones objetivo.

En cuanto a la selección de datos, se eligió un conjunto relevante del sector logístico, garantizando que ofreciera la información necesaria para desarrollar los modelos. Esta elección permitió abordar los diferentes problemas desde diversas perspectivas, asegurando que los datos fueran pertinentes y aplicables.

En el contexto logístico, se entrenaron cuatro modelos de regresión con el objetivo de predecir con precisión los tiempos de transporte. El modelo de bosque aleatorio demostró ser el más preciso, con un error cuadrático medio de 0.275, mientras que la regresión lineal mostró un rendimiento menor, con un error de 0.440. Este análisis permitió identificar las técnicas más adecuadas para este tipo de predicciones en el entorno logístico.

Para mejorar la eficiencia operativa y reducir los costos, se investigaron y desarrollaron estrategias de optimización de rutas de distribución. La implementación de estas estrategias se llevó a cabo utilizando la librería Google OR-Tools, que facilitó la identificación de rutas óptimas basadas en los datos recopilados, mostrando así la utilidad de las herramientas de optimización en la gestión logística.

Finalmente, se propuso un marco práctico y adaptable para la implementación eficiente de soluciones de aprendizaje automático en producción. Este marco combina conocimientos teóricos y prácticos para proporcionar un conjunto de herramientas que faciliten la integración de modelos de aprendizaje automático en entornos empresariales, abordando áreas clave como la evaluación, la estrategia y la acción.

La plataforma desarrollada con Kivy ofrece una interfaz gráfica interactiva que facilita la visualización y gestión de los modelos creados. A través de esta interfaz, los usuarios pueden acceder a representaciones visuales de predicciones y resultados, como imágenes y tablas que detallan rutas óptimas y tiempos de transporte estimados. La plataforma, que se ejecuta en una instancia de Linux en AWS Lightsail, se integra eficazmente con la arquitectura en la nube, permitiendo una gestión eficiente y accesible de los modelos de aprendizaje automático en producción.

REFERENCIAS

[1] A. Rushton, P. Croucher, and P. Baker, *The Handbook of Logistics and Distribution Management: Understanding the Supply Chain*. Kogan Page Publishers, 2022. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] M. S. Soria Uriarte and H. M. Lema Meneces, "Diseño de un modelo de gestión para la logística y distribución de productos de consumo masivo de una empresa ecuatoriana," Bachelor's thesis, 2019.

[3] A. N. Aspilcueta Arias, "Gestión de aprovisionamiento adaptado a Mypes comercializadoras de productos de consumo masivo para incrementar las utilidades operacionales: caso 'Corporación Chrimimepa SRL,'" Bachelor's thesis, 2019.

[4] V. P. Vergara Pinedo, "Desarrollo de Inteligencia de Negocios con un modelo de Machine Learning para la gestión de la flota vehicular," Bachelor's thesis, 2021.

[5] G. P. Negrete Godoy, "Diseño de un modelo Machine Learning para mejorar la venta de rutas diarias en la Startup de logística Wareclouds," Bachelor's thesis, 2022.

[6] J. L. Chavez Arzapalo, "Optimización de la distribución de carga en la empresa RPB SCM Logistics en Lima, Perú, utilizando Machine Learning: Un análisis de costos y tiempos de entrega," Bachelor's thesis, 2023.

[7] T. Kalganova and I. Dzalbs, "Supply Chain Logistics Problem Dataset," <https://doi.org/10.17633/rd.brunel.7558679.v2>, 2019.

[8] M. Akbari and T. N. A. Do, "A systematic review of machine learning in logistics and supply chain management: current trends and future directions," *Benchmarking: An International Journal*, vol. 28, no. 10, pp. 2977-3005, 2021.

[9] J. Tang, L. Zheng, C. Han, W. Yin, Y. Zhang, Y. Zou, and H. Huang, "Statistical and machine-learning methods for clearance time prediction of road incidents: A methodology review," *Analytic Methods in Accident Research*, vol. 27, p. 100123, 2020.

[10] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, *An Introduction to Statistical Learning: With Applications in Python*, pp. 69-134. Springer, 2023.

[11] J. Suzuki, *Statistical Learning with Math and R: 100 Exercises for Building Logic*, pp. 117-146. Springer, 2020.

[12] A. Liaw, M. Wiener, et al., "Classification and regression by randomForest," *R News*, vol. 2, no. 3, pp. 18-22, 2002.

[13] S. Fafalios, P. Charonyktakis, and I. Tsamardinos, "Gradient Boosting Trees," *Gnosis Data Analysis PC*, pp. 1-3, 2020.

[14] Y. Matsuo, et al., "Deep learning, reinforcement learning, and world models," *Neural Networks*, vol. 152, pp. 267-275, 2022.

[15] M. K. Satheesh and K. V. R. Kumar, "Addressing the Utilization of Popular Regression Models in Business Applications," in *Machine Learning for Business Analytics*, pp. 29-43. Productivity Press, 2022.

[16] X. Song and J. Dong, "A dynamic pricing and inventory model for retail chains with online and offline sales channels," *European Journal of Operational Research*, vol. 290, no. 3, pp. 915-927, 2021.

[17] M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, V. A. Kamaev, et al., "A survey of forecast error measures," *World Applied Sciences Journal*, vol. 24, no. 24, pp. 171-176, 2013.

[18] J. Chen, W. Liao, and C. Yu, "Route optimization for cold chain logistics of front warehouses based on traffic congestion and carbon emission," *Computers & Industrial Engineering*, vol. 161, p. 107663, 2021.

[19] W. Liu, "Route optimization for last-mile distribution of rural E-commerce logistics based on ant colony optimization," *IEEE Access*, vol. 8, pp. 12179-12187, 2020.

[20] M. Noureddine and M. Ristic, "Route planning for hazardous materials transportation: Multicriteria decision making approach," *Decision Making: Applications in Management and Engineering*, vol. 2, no. 1, pp. 66-85, 2019.

[21] A. Javaid, "Understanding Dijkstra's algorithm," *Available at SSRN 2340905*, 2013.

[22]